

(FILE 'HOME' ENTERED AT 09:36:26 ON 07 NOV 2003)

FILE 'REGISTRY' ENTERED AT 09:36:41 ON 07 NOV 2003

L1 2 S CGGTTTAATGGCTTGTGTGCT/SQSN  
L2 2 S L1 AND 15-30/SQL

FILE 'CAPLUS' ENTERED AT 09:38:19 ON 07 NOV 2003

L3 1 S L2  
L4 0 S L1 AND 15-100/SQL

FILE 'REGISTRY' ENTERED AT 09:39:48 ON 07 NOV 2003

L5 2 S L1 AND 15-100/SQL  
L6 8 S ATGCCATTAAACCGGTGGC/SQSN  
L7 2 S L6 AND 15-100/SQL

FILE 'CAPLUS' ENTERED AT 09:42:23 ON 07 NOV 2003

L8 6 S L6

FILE 'REGISTRY' ENTERED AT 09:46:05 ON 07 NOV 2003

FILE 'STNGUIDE' ENTERED AT 09:46:40 ON 07 NOV 2003

L9 QUE CGGTTTAATGGCTTGTGTGCT| AGCACAACAAGCCATTAAACCG| ATGCCATTAA

FILE 'REGISTRY' ENTERED AT 09:53:36 ON 07 NOV 2003

L10 14 S L9  
L11 6 S L10 AND SQL<101

FILE 'CAPLUS' ENTERED AT 09:54:58 ON 07 NOV 2003

L12 1 S L11  
L13 6 S L10

=>

ANSWER 4 OF 8    REGISTRY    COPYRIGHT 2003 ACS on STN

SEQ 27201 gcagtcctgat gccattaaac cggtaggctat tttccttgat ggctttgcgt

== =====

HITS AT:    27209-27227

**\*\*RELATED SEQUENCES AVAILABLE WITH SEQLINK\*\***

RN 405050-93-3    REGISTRY

CN DNA (Escherichia coli strain O157:H7 clone OZID\_244 fragment) (9CI) (CA  
INDEX NAME)

OTHER NAMES:

CN 242: PN: US6365723 SEQID: 242 claimed DNA

L13 ANSWER 1 OF 6 CAPLUS COPYRIGHT 2003 ACS on STN  
 ACCESSION NUMBER: 2002:937301 CAPLUS  
 DOCUMENT NUMBER: 138:20541  
 TITLE: DNA and protein sequences of enterohemorrhagic E. coli  
 0157:H7 specific proteins and their used in diagnosis  
 and therapeutics  
 INVENTOR(S): Hayashi, Hideo; Shinagawa, Hideo; Makino, Kozo;  
 Hayashi, Tetsuya; Onishi, Makoto; Hattori, Shohei;  
 Kurokawa, Akira  
 PATENT ASSIGNEE(S): Tsukuba University, Japan  
 SOURCE: Jpn. Kokai Tokkyo Koho, 2067 pp.  
 CODEN: JKXXAF  
 DOCUMENT TYPE: Patent  
 LANGUAGE: Japanese  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

	PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
	JP 2002355074	A2	20021210	JP 2002-15959	20020124
PRIORITY APPLN. INFO.:				JP 2001-112010	A 20010124
AB	This invention provides DNA and protein sequences of enterohemorrhagic E. coli O157:H7 specific proteins. The proteins are specifically present in E. coli O157:H7 rather than in non-pathogenic E. coli K-12. The O157:H7 specific DNA and protein can be used in diagnosis and treatment of digestive track hemorrhage.				

IT	478201-74-0	478204-20-5	478205-30-0	478206-19-8	478206-71-2
	478207-40-8	478207-99-7	478208-63-8	478209-11-9	478209-77-7
	478210-23-0	478210-82-1	478211-21-1	<b>478211-56-2</b>	
	478211-72-2	478211-98-2	478212-03-2	478212-30-5	478212-65-6
	478212-92-9	478213-26-2	478213-47-7	478213-75-1	478213-93-3
	478214-09-4	478214-23-2	478214-37-8	478214-44-7	478214-54-9
	478214-62-9	478214-71-0	478214-81-2	478214-91-4	478214-96-9
	478215-04-2	478215-10-0	478215-16-6	478215-24-6	478215-31-5
	478215-38-2	478215-41-7	478215-48-4	478215-55-3	478215-64-4
	478215-71-3	478215-78-0	478215-85-9	478215-88-2	478215-95-1
	478215-99-5	478216-01-2	478216-06-7	478216-15-8	478216-17-0
	478216-22-7	478216-27-2	478216-29-4	478216-34-1	478216-36-3
	478216-38-5	478216-39-6	478216-44-3	478216-48-7	478216-54-5
	478216-59-0	478216-64-7	478216-67-0	478216-72-7	478216-74-9
	478216-76-1	478216-78-3	478216-80-7	478216-82-9	478216-85-2
	478216-90-9	478216-93-2	478216-96-5	478216-98-7	478217-03-7
	478217-04-8	478217-05-9	478217-08-2	478217-12-8	478217-15-1
	478217-18-4	478217-21-9	478217-22-0	478217-23-1	478217-25-3
	478217-30-0	478217-33-3	478217-35-5	478217-37-7	478217-40-2
	478217-42-4	478217-44-6	478217-45-7	478217-47-9	478217-50-4
	478217-53-7	478217-54-8	478217-55-9	478217-57-1	478217-58-2
	478217-60-6	478217-61-7	478217-64-0	478217-65-1	478217-66-2
	478217-67-3	478217-69-5	478217-71-9	478217-73-1	478217-75-3
	478217-77-5	478217-78-6	478217-79-7	478217-81-1	478217-82-2
	478217-84-4	478217-86-6	478217-89-9	478217-91-3	478217-92-4
	478217-94-6	478217-97-9	478217-99-1	478218-00-7	478218-03-0
	478218-05-2	478218-08-5	478218-09-6	478218-10-9	478218-13-2
	478218-14-3	478218-15-4	478218-16-5	478218-17-6	478218-18-7
	478218-19-8	478218-20-1	478218-21-2	478218-22-3	

RL: BSU (Biological study, unclassified); DGN (Diagnostic use); PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES (Uses) (nucleotide sequence; DNA and protein sequences of enterohemorrhagic E. coli O157:H7 specific proteins and their used in diagnosis and therapeutics)

L13 ANSWER 2 OF 6 CAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2002:251953 CAPLUS  
DOCUMENT NUMBER: 136:258372  
TITLE: DNA sequences unique to the pathogenic O157:H7 strain  
of Escherichia coli  
INVENTOR(S): Blattner, Frederick R.; Burland, Valerie; Perna,  
Nicole T.; Plunkett, Guy; Welch, Rod  
PATENT ASSIGNEE(S): Wisconsin Alumni Research Foundation, USA  
SOURCE: U.S., 26 pp.  
CODEN: USXXAM  
DOCUMENT TYPE: Patent  
LANGUAGE: English  
FAMILY ACC. NUM. COUNT: 1  
PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
US 6365723	B1	20020402	US 1999-453702	19991203
US 2003023075	A1	20030130	US 2002-114170	20020401

PRIORITY APPLN. INFO.: US 1998-110955P P 19981204  
US 1999-453702 A1 19991203

AB The entire genome of pathogenic Escherichia coli strain O157:H7 has been sequenced. All of the genomic DNA sequences present in O157 and absent in the previously sequenced laboratory strain K12 are provided. These sequences correspond to 255 islands of O157 DNA. This sequence information is needed for comprehensive efforts at detection, diagnosis, prophylaxis, and therapeutic approaches to infections caused by the organism.

REFERENCE COUNT: 4 THERE ARE 4 CITED REFERENCES AVAILABLE FOR THIS RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

IT 405050-88-6	405050-89-7	405050-90-0	405050-91-1	405050-92-2
<b>405050-93-3</b>	405050-94-4	405050-95-5	405050-96-6	
405050-97-7	405050-98-8	405050-99-9	405051-00-5	405051-01-6
405051-02-7	405051-03-8	405051-04-9	405051-05-0	405051-06-1
405051-07-2	405051-08-3	405051-09-4	405051-10-7	405051-11-8
405051-12-9	405051-14-1			

RL: BSU (Biological study, unclassified); PRP (Properties); THU  
(Therapeutic use); BIOL (Biological study); USES (Uses)  
(nucleotide sequence; DNA sequences unique to the pathogenic O157:H7 strain of Escherichia coli)

L13 ANSWER 3 OF 6 CAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2001:209011 CAPLUS  
DOCUMENT NUMBER: 137:196299  
TITLE: Genome sequence of enterohaemorrhagic Escherichia coli  
O157:H7. [Erratum to document cited in CA134:232542]  
AUTHOR(S): Perna, Nicole T.; Plunkett, Guy, III; Burtand,  
Valerie; Mau, Bob; Glasner, Jeremy D.; Rose, Debra J.;  
Mayhew, George F.; Evans, Peter S.; Gregor, Jason;  
Kirkpatrick, Heather A.; Postal, Gyorgy; Hackett,  
Jeremiah; Klink, Sara; Boutin, Adam; Shao, Ying;  
Miller, Leslie; Grotbeck, Erik J.; Davis, N. Wayne;  
Lim, Alex; Dimalanta, Eileen T.; Potamousis,  
Konstantinos D.; Apodaca, Jennifer; Anantharaman,  
Thomas S.; Lin, Jleyl; Yen, Galex; Schwartz, Dvauid C.;  
Welch, Rodney A.; Blattner, Frederick R.  
CORPORATE SOURCE: Genome Center of Wisconsin, Department of Animal  
Health and Biomedical Sciences, Laboratory of  
Genetics, Department of Chemistry, Department of  
Biostatistics, and Department of Medical Microbiology  
and Immunology, University of Wisconsin, Madison, WI,  
53706, USA  
SOURCE: Nature (London, United Kingdom) (2001), 410(6825), 240  
CODEN: NATUAS; ISSN: 0028-0836

PUBLISHER: Nature Publishing Group  
DOCUMENT TYPE: Journal  
LANGUAGE: English

AB The correct GenBank accession number for the annotated sequence is AE005174.  
IT 318232-68-7, GenBank AE005653 318232-69-8, GenBank AE005654  
318232-70-1, GenBank AE005655 318232-71-2, GenBank AE005656  
318232-72-3, GenBank AE005657 318232-73-4, GenBank AE005658  
318232-74-5, GenBank AE005659 318232-75-6, GenBank AE005660  
**318232-76-7**, GenBank AE005661 318232-77-8, GenBank AE005662  
318232-78-9, GenBank AE005663 318232-79-0, GenBank AE005664  
318232-80-3, GenBank AE005665 318232-81-4, GenBank AE005666  
318232-82-5, GenBank AE005667 318232-83-6, GenBank AE005668  
318232-84-7, GenBank AE005669 318232-85-8, GenBank AE005670  
318232-86-9, GenBank AE005671  
RL: BSU (Biological study, unclassified); PRP (Properties); BIOL  
(Biological study)  
(nucleotide sequence; genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 (Erratum))

L13 ANSWER 4 OF 6 CAPLUS COPYRIGHT 2003 ACS on STN  
ACCESSION NUMBER: 2001:184271 CAPLUS  
DOCUMENT NUMBER: 134:217892  
TITLE: Complete genome sequence of enterohemorrhagic  
*Escherichia coli* O157:H7 and genomic comparison with a  
laboratory strain K-12  
AUTHOR(S): Hayashi, Tetsuya; Makino, Kozo; Ohnishi, Makoto;  
Kurokawa, Ken; Ishii, Kazuo; Yokoyama, Katsushi; Han,  
Chang-Gyun; Ohtsubo, Eiichi; Nakayama, Keisuke;  
Murata, Takahiro; Tanaka, Masashi; Tobe, Toru; Iida,  
Tetsuya; Takami, Hideto; Honda, Takeshi; Sasakawa,  
Chihiro; Ogasawara, Naotake; Yasunaga, Teruo; Kuhara,  
Satoru; Shiba, Tadayoshi; Hattori, Masahira;  
Shinagawa, Hideo  
CORPORATE SOURCE: Department of Microbiology, Miyazaki Medical College,  
Miyazaki, 899-1692, Japan  
SOURCE: DNA Research (2001), 8(1), 11-22  
CODEN: DARSE8; ISSN: 1340-2838  
PUBLISHER: Universal Academy Press  
DOCUMENT TYPE: Journal  
LANGUAGE: English

AB *Escherichia coli* O157:H7 is a major food-borne infectious pathogen that  
causes diarrhea, hemorrhagic colitis, and hemolytic uremic syndrome. The  
complete chromosome sequence of an O157:H7 strain isolated from the Sakai  
outbreak is reported, and the results compared with the genome of a benign  
laboratory strain, K-12 MG1655. The chromosome is 5.5 Mb in size, 859 Kb  
larger  
than that of K-12. A 4.1-Mb sequence highly conserved between the two  
strains is identified, which may represent the fundamental backbone of the  
*E. coli* chromosome. The remaining 1.4-Mb sequence comprises of  
O157:H7-specific sequences, most of which are horizontally transferred  
foreign DNAs. The predominant roles of bacteriophages in the emergence of  
O157:H7 is evident by the presence of 24 prophages and prophage-like  
elements that occupy more than half of the O157:H7-specific sequences.  
The O157:H7 chromosome encodes 1632 proteins and 20 tRNAs that are not  
present in K-12. Among these, at least 131 proteins are assumed to have  
virulence-related functions. Genome-wide codon usage anal. suggested that  
the O157:H7-specific tRNAs are involved in the efficient expression of the  
strain-specific genes. A complete set of the genes specific to O157:H7  
presented here sheds new insight into the pathogenicity and the physiol.  
of O157:H7, and will open a way to fully understand the mol. mechanisms  
underlying the O157:H7 infection.

REFERENCE COUNT: 62 THERE ARE 62 CITED REFERENCES AVAILABLE FOR THIS  
RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

IT 328224-91-5, GenBank AP002550 328224-92-6, GenBank AP002551

328224-93-7, GenBank AP002552      328224-94-8, GenBank AP002553  
 328224-95-9, GenBank AP002554      328224-96-0, GenBank AP002555  
 328224-97-1, GenBank AP002556      328224-98-2, GenBank AP002557  
 328224-99-3, GenBank AP002558      328225-00-9, GenBank AP002559  
 328225-01-0, GenBank AP002560      328225-02-1, GenBank AP002561  
 328225-03-2, GenBank AP002562      328225-04-3, GenBank AP002563  
 328225-05-4, GenBank AP002564      328225-06-5, GenBank AP002565  
 328225-07-6, GenBank AP002566      328225-08-7, GenBank AP002567  
 328225-09-8, GenBank AP002568      328225-10-1, GenBank AP002569  
 RL: BSU (Biological study, unclassified); PRP (Properties); BIOL  
 (Biological study)  
 (nucleotide sequence; complete genome sequence of enterohemorrhagic  
 Escherichia coli O157:H7 and genomic comparison with a laboratory strain  
 K-12)

L13 ANSWER 5 OF 6 CAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2001:98372 CAPLUS

DOCUMENT NUMBER: 134:232542

TITLE: Genome sequence of enterohaemorrhagic Escherichia coli  
 O157:H7

AUTHOR(S): Perna, Nicole T.; Plunkett, Guy, III; Burland,  
 Valerie; Mau, Bob; Glasner, Jeremy D.; Rose, Debra J.;  
 Mayhew, George F.; Evans, Peter S.; Gregor, Jason;  
 Kirkpatrick, Heather A.; Posfai, Gyorgy; Hackett,  
 Jeremiah; Klink, Sara; Boutin, Adam; Shao, Ying;  
 Miller, Leslie; Grotbeck, Erik J.; Davis, N. Wayne;  
 Lim, Alex; Dimalanta, Eileen T.; Potamousis,  
 Konstantinos D.; Apodaca, Jennifer; Anantharaman,  
 Thomas S.; Lin, Jieyi; Yen, Glaex; Schwartz, David C.;  
 Welch, Rodney A.; Blattner, Frederick R.

CORPORATE SOURCE: Genome Center of Wisconsin, Department of Animal  
 Health and Biomedical Sciences, Laboratory of  
 Genetics, Department of Chemistry, Department of  
 Biostatistics, and Department of Medical Microbiology  
 and Immunology, University of Wisconsin, Madison, WI,  
 53706, USA

SOURCE: Nature (London) (2001), 409(6819), 529-533

CODEN: NATUAS; ISSN: 0028-0836

PUBLISHER: Nature Publishing Group

DOCUMENT TYPE: Journal

LANGUAGE: English

AB The bacterium Escherichia coli O157:H7 is a worldwide threat to public  
 health and has been implicated in many outbreaks of hemorrhagic colitis,  
 some of which included fatalities caused by hemolytic uremic syndrome.  
 Close to 75,000 cases of O157:H7 infection are now estimated to occur annually  
 in the United States. The severity of disease, the lack of effective  
 treatment and the potential for large-scale outbreaks from contaminated  
 food supplies have propelled intensive research on the pathogenesis and  
 detection of E. coli O157:H7. The genome of E. coli O157:H7 was sequenced  
 to identify candidate genes responsible for pathogenesis, to develop  
 better methods of strain detection and to advance our understanding of the  
 evolution of E. coli, through comparison with the genome of the  
 non-pathogenic laboratory strain E. coli K-12. Lateral gene transfer found to  
 be far more extensive than previously anticipated. In fact, 1387 new  
 genes encoded in strain-specific clusters of diverse sizes were found in  
 O157:H7. These include candidate virulence factors, alternative metabolic  
 capacities, several prophages, and other new functions - all of which  
 could be targets for surveillance.

REFERENCE COUNT: 30 THERE ARE 30 CITED REFERENCES AVAILABLE FOR THIS  
 RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

IT 318232-68-7, GenBank AE005653      318232-69-8, GenBank AE005654  
 318232-70-1, GenBank AE005655      318232-71-2, GenBank AE005656  
 318232-72-3, GenBank AE005657      318232-73-4, GenBank AE005658  
 318232-74-5, GenBank AE005659      318232-75-6, GenBank AE005660

318232-76-7, GenBank AE005661      318232-77-8, GenBank AE005662  
 318232-78-9, GenBank AE005663      318232-79-0, GenBank AE005664  
 318232-80-3, GenBank AE005665      318232-81-4, GenBank AE005666  
 318232-82-5, GenBank AE005667      318232-83-6, GenBank AE005668  
 318232-84-7, GenBank AE005669      318232-85-8, GenBank AE005670  
 318232-86-9, GenBank AE005671  
 RL: BSU (Biological study, unclassified); PRP (Properties); BIOL  
 (Biological study)  
 (nucleotide sequence; genome sequence of enterohemorrhagic Escherichia  
 coli O157:H7)

L13 ANSWER 6 OF 6 CAPLUS COPYRIGHT 2003 ACS on STN

ACCESSION NUMBER: 2000:900842 CAPLUS  
 DOCUMENT NUMBER: 134:52236  
 TITLE: A DNA marker specific for Escherichia coli serotypes  
 O157:H7; O157:NM and O55:H7 and its diagnostic use  
 INVENTOR(S): Chen, Shu; Xu, Renlin; Li, Jiping  
 PATENT ASSIGNEE(S): University of Guelph, Can.  
 SOURCE: PCT Int. Appl., 40 pp.  
 CODEN: PIXXD2  
 DOCUMENT TYPE: Patent  
 LANGUAGE: English  
 FAMILY ACC. NUM. COUNT: 1  
 PATENT INFORMATION:

PATENT NO.	KIND	DATE	APPLICATION NO.	DATE
WO 2000077247	A1	20001221	WO 2000-CA716	20000614
W: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU,				
CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL,				
IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA,				
MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI,				
SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW, AM,				
AZ, BY, KG, KZ, MD, RU, TJ, TM				
RW: GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW, AT, BE, CH, CY,				
DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, BF, BJ,				
CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG				

PRIORITY APPLN. INFO.: US 1999-139260P P 19990615

AB A novel DNA marker FAFLP (fluorescent amplified fragment length  
 polymorphism) specific for E. coli serotypes O157:H7; O157:NM and O55:H7  
 is disclosed. The isolation of the marker allows the development of  
 diagnostic assays that can be used to detect the serotypes in the sample.  
 In particular, the marker can be used to prepare nucleic acid primers and  
 nucleotide probes based on the sequence of the marker.

REFERENCE COUNT: 12 THERE ARE 12 CITED REFERENCES AVAILABLE FOR THIS  
 RECORD. ALL CITATIONS AVAILABLE IN THE RE FORMAT

IT 313725-55-2 313725-56-3  
 RL: PRP (Properties); THU (Therapeutic use); BIOL (Biological study); USES  
 (Uses)

(nucleotide sequence; a DNA marker specific for Escherichia coli  
 serotypes O157:H7, O157:NM and O55:H7 and diagnostic use)

IT 313725-50-7 313725-51-8 313725-52-9  
 313725-54-1  
 RL: ARG (Analytical reagent use); PRP (Properties); ANST (Analytical  
 study); USES (Uses)  
 (nucleotide sequences of primers or probes; a DNA marker specific for  
 Escherichia coli serotypes O157:H7, O157:NM and O55:H7 and diagnostic  
 use)

TITLE Direct Submission  
 JOURNAL Submitted (26-JUN-2000) Ken Kurokawa, Osaka University, Genome  
 Information Research Center; 3-1, Yamadaoka, Suita, Osaka 565-0871,  
 Japan (E-mail:ken@gen-info.osaka-u.ac.jp,  
 URL:http://www.gen-info.osaka-u.ac.jp/, Tel:81-6-6879-8365,  
 Fax:81-6-6879-2047)  
 COMMENT genome project.  
 FEATURES Location/Qualifiers  
     source 1. .168764  
         /organism="Escherichia coli O157:H7"  
         /mol\_type="genomic DNA"  
         /strain="O157:H7"  
         /sub\_strain="RIMD 0509952"  
         /db\_xref="taxon:83334"  
     gene 79. .243  
         /gene="ECs5227"  
     CDS 79. .243  
         /gene="ECs5227"  
         /note="identical to B4250\_ECOLI gi|1790699 (Conserved in  
         E.coli K-12)"  
         /codon\_start=1  
         /evidence=not\_experimental  
         /transl\_table=11  
         /product="hypothetical protein"  
         /protein\_id="BAB38650.1"  
         /db\_xref="GI:13364705"  
         /translation="MLSVLLKLLIILVLTTPRASTRLLAVRLGYLAVYSMNMSVRKLFRL  
         LSIFFVMIVR"  
     gene 315. .569  
         /gene="ECs5228"  
     CDS 315. .569  
         /gene="ECs5228"  
         /note="similar to YJGJ\_ECOLI gi|1790700 percent identity  
         92 in 84 aa (Conserved in E.coli K-12)"  
         /codon\_start=1  
         /evidence=not\_experimental  
         /transl\_table=11  
         /product="hypothetical protein"  
         /protein\_id="BAB38651.1"  
         /db\_xref="GI:13364706"  
         /translation="MVLEGIHSHDPQARDIAVQYYHAAETAIYDYIARLHPQSAQC  
         VTD FMSTVMSGLSAKAREGHSLEQLCATAALAGEAIAIKTILKE"  
     gene 705. .1166  
         /gene="ECs5229"  
     CDS 705. .1166  
         /gene="ECs5229"  
         /note="similar to YJGK\_ECOLI gi|1790701 percent identity  
         98 in 153 aa (Conserved in E.coli K-12)"  
         /codon\_start=1  
         /evidence=not\_experimental  
         /transl\_table=11  
         /product="hypothetical protein"  
         /protein\_id="BAB38652.1"  
         /db\_xref="GI:13364707"  
         /translation="MEKMIIGNIHNLPWLPQELRQAIEHIKAHVTAETPKGKHDIEG  
         NRLFYLISEDMTEPYEARAEYHARYLDIQIVLKGQEGMTFSTQPA GTPD TDWLADKD  
         IAFLPEGVDEKTVILNEGDFVVFYPGEVHKPLCAVGAPAQVRKAVVKMLMA"



gene 1289. .3061  
 /gene="ECs5230"  
 CDS 1289. .3061  
 /gene="ECs5230"  
 /note="similar to part of (13-616 in 616 aa) YJGL\_ECOLI gi|1790702 percent identity 82 in 590 aa"  
 /codon\_start=1  
 /evidence=not\_experimental  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="BAB38653.1"  
 /db\_xref="GI:13364708"  
 /translation="MSKISDLNYSQHITLADNFKQKSEVLNTWRVGMNMFARNAGGQD  
 NTRNILNPKTFLEFLVKIFTLGYVDFSKRSNEAGRNMMAHIESSSYIKNNDGSEIMKF  
 VMNPEGERADSPKVIIEISLSTITTMGTRQGHTAIIFFQPDGSTNRYERKSFERKDE  
 SSLHLITNKVLACYQREANKEIARLLNNHQKLNNLQKLNNLQKLNNLQKLNNIQLN  
 IQELNNSQELNNSQELNNSQELNNSQDLKNSQVSCCKGSVDFTITDLLEKSLNNALLAI  
 RNEHLLMPHVCSEISISYLLGENGILEEIDKLYELNDHGIDNDKEGNNEINDIMINLS  
 HILIESLDDAKVNLTPVIHSMMLTFLELPYNNNDVKILEWCFNKSMMQYFDDSAKIEHAC  
 SVINHINFRRDQSKVAETLFFNLDKEPYKNSPELQELIWKKLVVYVNDFNLSNREKTY  
 LIQRIFNNVESLFNKVPVSILVNDIFMNDFFMKNTMINWYFPRLLKSYEDEKIYFDK  
 LGYNFNNKESNEEIMKNQPKDVIEEKLNNELKLRFRMMQTILKSEVNVSPFIDQQRNLN  
 TLNPPENLRIAIEKFGWKKKTITA"  
 gene complement(3118. .4122)  
 /gene="ECs5231"  
 CDS complement(3118. .4122)  
 /gene="ECs5231"  
 /note="similar to ARGF\_ECOLI gi|1786469 percent identity  
 87 in 334 aa, similar to ARG1\_ECOLI gi|1790703 percent  
 identity 99 in 334 aa (Conserved in E.coli K-12)"  
 /codon\_start=1  
 /evidence=not\_experimental  
 /transl\_table=11  
 /product="ornithine carbamoyltransferase 1"  
 /protein\_id="BAB38654.1"  
 /db\_xref="GI:13364709"  
 /translation="MSGFYHKHFLKLLDFTPAELNSLLQLAAKLKADKKSGKEEAKLT  
 GKNIALIFEKDSTRTRCSFEVAAYDQGARVTYLGPSPGSQIGHKESIKDTARVLGRMYD  
 GIQYRGYQQEIVETLAEYAGVPVWNGLTNEFHPTQLLADLLTMQEHLPGKPFNEMTLV  
 YAGDARNNMGNSMLEAAALTGLDLRLVAPQACWPEAALVTECRALAQQNGGNITLTED  
 VAKGVEGADFIYTDVWVSMGEAKEKWAERIALLRDYQVNSKMMQLTGNPEVKFLHCLP  
 AFHDDQTTLGKKMAEEFGLHGGMEVTDDEVFESAASIVFDQAENRMHTIKAVMVATLSK  
 "  
 gene 4284. .4700  
 /gene="ECs5232"  
 CDS 4284. .4700  
 /gene="ECs5232"  
 /note="similar to YJGD\_ECOLI gi|1790704 percent identity  
 100 in 138 aa (Conserved in E.coli K-12)"  
 /codon\_start=1  
 /evidence=not\_experimental  
 /transl\_table=11  
 /product="hypothetical protein"  
 /protein\_id="BAB38655.1"  
 /db\_xref="GI:13364710"  
 /translation="MANPEQLEEQREETRLIIIEELLEDDGSDPDALYTIEHHSADDLE  
 TLEKAAVEAFKLGYEVTDPPEELEVEDGDIVICCDILSECALNADLIDAQVEQLMTLAE

```

gene      KFDVEYDGWGTYFEDPNGEDGDDEDFVDEDDDGVRH"
          complement(4746..5039)
          /gene="ECs5233"
CDS       complement(4746..5039)
          /gene="ECs5233"
          /note="similar to B4256_ECOLI gi|1790705 percent identity
          93 in 97 aa (Conserved in E.coli K-12)"
          /codon_start=1
          /evidence=not_experimental
          /transl_table=11
          /product="hypothetical protein"
          /protein_id="BAB38656.1"
          /db_xref="GI:13364711"
          /translation="MVGGGGIAPLAGSESDICELQKMYFLPAIRGKGLAKKLALKAME
          EAREMGFKRCYLETTAFLKEAIGLYEHLGFQHIDYALGCTGHVDCEVRMLREL"
gene      5442..6638
          /gene="ECs5234"
CDS       5442..6638
          /gene="ECs5234"
          /note="similar to part (5-393 in 393 aa)of YJGN_ECOLI
          gi|2367367 percent identity 96 in 389 aa (Conserved in
          E.coli K-12)"
          /codon_start=1
          /evidence=not_experimental
          /transl_table=11
          /product="hypothetical protein"
          /protein_id="BAB38657.1"
          /db_xref="GI:13364712"
          /translation="MAQVINEMDVPSHSFVFHGTGERYFLICVVNVLLTIITLGIYLP
          WALMKCKRYLYANMEVNGQRFSYGITGGNVFFSCLVFFVFFYFAILMTVSADMP LIGCV
          LTLSLLVLLIFMAAKGLRYQALMTSLNGVRFSFNCMSMKGVWWVTFFLPILMAIGMTV

```

```

Query Match          99.3%;   Score 1572;   DB 1;   Length 168764;
Best Local Similarity 99.9%;   Pred. No. 0;
Matches 1583;   Conservative    0;   Mismatches    0;   Indels    1;   Gaps    1;

```

```

Qy      1 CTGCACCTTTTTTTGCTGTGCCTTTGGGATCCGCGCGAATGCGCCCAGAAAACCATCAGA 60
          |||
Db      58317 CTGCACCTTTTTTTGCTGTGCCTTTGGGATCCGCGCGAATGCGCCCAGAAAACCATCAGA 58258
          |||

Qy      61 AAGTGCCACCTTGGGCGTTTGCACCGTCATTTGTACCTCCGGACAGATGTTGCAATGTA 120
          |||
Db      58257 AAGTGCCACCTTGGGCGTTTGCACCGTCATTTGTACCTCCGGACAGATGTTGCAATGTA 58198
          |||

Qy      121 TTTTCGTCAATTGGCCCCGAATGCAACGTGCCAACCCCTGGTTGTGAATAATGGAATAAAAT 180
          |||
Db      58197 TTTTCGTCAATTGGCCCCGAATGCAACGTGCCAACCCCTGGTTGTGAATAATGGAATAAAAT 58138
          |||

Qy      181 TCTTGATTATCAATAATCAATGCCCATTTTGTGATGGCCAGGCGAGATCGGCTTCCGCA 240
          |||
Db      58137 TCTTGATTATCAATAATCAATGCCCATTTTGTGATGGCCAGGCGAGATCGGCTTCCGCA 58078
          |||

Qy      241 ATAATTTACCTTCATCATTTTGCAATTCGTAGCCGAGAGTCGGCGCAGAGAGAGAATGA 300
          |||
Db      58077 ATAATTTACCTTCATCATTTTGCAATTCGTAGCCGAGAGTCGGCGCAGAGAGAGAATGA 58018
          |||

Qy      301 GACTGAAGCAACGCGATCTCTTCAGCGCTTAATAAGCCAAACTCAATAATCTCAGCCCAA 360

```

Db	58017	 GACTGAAGCAACGCGATCTCTTCAGCGCTTAATAAGCCAAACTCAATAATCTCAGCCCAA	57958
Qy	361	CTCTCATCAGTCTCTGGTACTGCCTCTGGCGTTGCCGAAACTTTAGTGTCAACGCTTTCC	420
Db	57957	 CTCTCATCAGTCTCTGGTACTGCCTCTGGCGTTGCCGAAACTTTAGTGTCAACGCTTTCC	57898
Qy	421	TGCGGTAAATGCACGGCTTTACGGCTGGTAAAGGTCATATCCGGCAGGAACTGTAACAGG	480
Db	57897	 TGCGGTAAATGCACGGCTTTACGGCTGGTAAAGGTCATATCCGGCAGGAACTGTAACAGG	57838
Qy	481	TTTACCAGTCGCCAGAAGCCATTGAGTCCCGCTTCATAACCGTCATCCTGCGTATAGCGA	540
Db	57837	 TTTACCAGTCGCCAGAAGCCATTGAGTCCCGCTTCATAACCGTCATCCTGCGTATAGCGA	57778
Qy	541	TCGTCAAAACAGATATGCAAACGCAGCCAGCTGCGCATTTGTGCGACGCTGGTCGTCGGT	600
Db	57777	 TCGTCAAAACAGATATGCAAACGCAGCCAGCTGCGCATTTGTGCGACGCTGGTCGTCGGT	57718
Qy	601	TTAATGGCTTGTGTGGTAACACCGAAGCCAGCTCAATAAATTGCTGCGATGAGTTACAG	660
Db	57717	 TTAATGGCTTGTGTGGTAACACCGAAGCCAGCTCAATAAATTGCTGCGATGAGTTACAG	57658
Qy	661	CTATCGAGTAAACCACCAAAGACAAACGGTTCGTCTGGCAAAAGCTCTGCTAGCCGCCAA	720
Db	57657	 CTATCGAGTAAACCACCAAAGACAAACGGTTCGTCTGGCAAAAGCTCTGCTAGCCGCCAA	57598
Qy	721	GCTGGCGCGTTTTCTGCATTTTCATAGGCATACTTTTGTTCGTTCCCGCGTCCTGTGAC	780
Db	57597	 GCTGGCGCGTTTTCTGCATTTTCATAGGCATACTTTTGTTCGTTCCCGCGTCCTGTGAC	57538
Qy	781	TTCTTCACATCCAACCACACCCAGGCATGTGCAGCAGCCATTCTCTGCCATAACTGCGTT	840
Db	57537	 TTCTTCACATCCAACCACACCCAGGCATGTGCAGCAGCCATTCTCTGCCATAACTGCGTT	57478
Qy	841	CTTCCGCCCCGATCGGCGAGATAATCCAGCAGCAGTGCAAAACTGTTGCGTTCCCGGAAC	900
Db	57477	 CTTCCGCCCCGATCGGCGAGATAATCCAGCAGCAGTGCAAAACTGTTGCGTTCCCGGAAC	57418
Qy	901	GACGCTTCAAGGGCCGCAAAGTTGGTATCATGAAACAGGTTATAAACTTAGGCTGTTTC	960
Db	57417	 GACGCTTCAAGGGCCGCAAAGTTGGTATCATGAAACAGGTTATAAACTTAGGCTGTTTC	57358
Qy	961	ATATCCGGGTTATGCCCAAGTCCCAGGACATCCTGAACATGTTTAAGCCCTGGCTCTTTG	1020
Db	57357	 ATATCCGGGTTATGCCCAAGTCCCAGGACATCCTGAACATGTTTAAGCCCTGGCTCTTTG	57298
Qy	1021	AGATCGG-CCACGTGACTGTCCACACCCAGAAATTACCGCTATCCCGAATCGCCTGGCGT	1079
Db	57297	 AGATCGGCCCCACGTGACTGTCCACACCCAGAAATTACCGCTATCCCGAATCGCCTGGCGT	57238
Qy	1080	TTTTGCACATCCTCTGCGACGCTATTTTTGTGGAACGCAAAGCCATCAAGGAAAATAGCC	1139
Db	57237	 TTTTGCACATCCTCTGCGACGCTATTTTTGTGGAACGCAAAGCCATCAAGGAAAATAGCC	57178
Qy	1140	ACCGGTTTAATGGCATCAGACTGCATAAGCGGATATAGCACGTAATCCGGGCGACAAGGG	1199

Db	57177	ACCGGTTTAAATGGCATCAGACTGCATAAGCGGATATAGCACGTAATCCGGGCGACAAGGG	57118
Qy	1200	ATCCCCACCGCCTCTTTAGCGCCTAAATCCACCTGAGCTTTAAGATGCCAGCTCATCATC	1259
Db	57117	ATCCCCACCGCCTCTTTAGCGCCTAAATCCACCTGAGCTTTAAGATGCCAGCTCATCATC	57058
Qy	1260	GGTTCGCTGCGGGTGTTAATAATCCACCCAGCGCCCTGATGCGCGTAACTGCGACTGACA	1319
Db	57057	GGTTCGCTGCGGGTGTTAATAATCCACCCAGCGCCCTGATGCGCGTAACTGCGACTGACA	56998
Qy	1320	ATCAGGTTTTTATTTTCTTGACAGGCAGCTAATAAACCGTTTTTCCAGCTCACTGCCCATC	1379
Db	56997	ATCAGGTTTTTATTTTCTTGACAGGCAGCTAATAAACCGTTTTTCCAGCTCACTGCCCATC	56938
Qy	1380	ATGGCTTCGAGGGAAATGTTTTTGATGGAATCAATTACCCGTATCGAATCTCTTGCCCTGC	1439
Db	56937	ATGGCTTCGAGGGAAATGTTTTTGATGGAATCAATTACCCGTATCGAATCTCTTGCCCTGC	56878
Qy	1440	AAAATTTTCGCCAGCAAGAGACGCGCCTGATCGCGGGAGACATACTTCATGCGCCACGG	1499
Db	56877	AAAATTTTCGCCAGCAAGAGACGCGCCTGATCGCGGGAGACATACTTCATGCGCCACGG	56818
Qy	1500	TCACGGTAAGCGTAAACGCAACGATAACAACCATCTTTATGCGTATCATTGTTGCAACTG	1559
Db	56817	TCACGGTAAGCGTAAACGCAACGATAACAACCATCTTTATGCGTATCATTGTTGCAACTG	56758
Qy	1560	CATTCCACTATTGCCTGATAGGCC	1583
Db	56757	CATTCCACTATTGCCTGATAGGCC	56734

```

RESULT 1
US-09-453-702B-242/c
; Sequence 242, Application US/09453702B
; Patent No. 6365723
;   GENERAL INFORMATION:
;       APPLICANT: Blattner, Frederick R.
;                   Burland, Valerie
;                   Perna, Nicole T.
;                   Plunkett, Guy
;                   Welch, Rod
;   TITLE OF INVENTION: No. 6365723el Sequences of E. coli O157
;   NUMBER OF SEQUENCES: 265
;   CORRESPONDENCE ADDRESS:
;       ADDRESSEE: Quarles & Brady
;       STREET: 1 South Pinckney Street
;       CITY: Madison
;       STATE: WI
;       COUNTRY: US
;       ZIP: 53701-2113
;   COMPUTER READABLE FORM:
;       MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage
;       COMPUTER: IBM PC compatible
;       OPERATING SYSTEM: PC-DOS/MS-DOS
;       SOFTWARE: Word Perfect 8.0
;   CURRENT APPLICATION DATA:
;       APPLICATION NUMBER: US/09/453,702B
;       FILING DATE: 03-Dec-1999
;       CLASSIFICATION: <Unknown>
;   PRIOR APPLICATION DATA:
;       APPLICATION NUMBER: 60/110,955
;       FILING DATE: 04-DEC-1998
;   ATTORNEY/AGENT INFORMATION:
;       NAME: Seay, Nicholas J.
;       REGISTRATION NUMBER: 27386
;       REFERENCE/DOCKET NUMBER: 960296.95017
;   TELECOMMUNICATION INFORMATION:
;       TELEPHONE: (608) 251-5000
;       TELEFAX: (608) 251-9166
;   INFORMATION FOR SEQ ID NO: 242:
;       SEQUENCE CHARACTERISTICS:
;           LENGTH: 31880
;           TYPE: nucleic acid
;           STRANDEDNESS: double
;           TOPOLOGY: linear
;       MOLECULE TYPE: DNA (genomic)
;       SEQUENCE DESCRIPTION: SEQ ID NO: 242:
US-09-453-702B-242

```

Query Match 100.0%; Score 96; DB 4; Length 31880;  
Best Local Similarity 100.0%; Pred. No. 2.7e-26;  
Matches 96; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 TATCCCGAATCGCCTGGCGTTTTTGCACATCCTCTGCGACGCTATTTTTGTGGAACGCAA 60  
 ||||||||||||||||||||||||||||||||||||||||||||||||||||  
 Db 27304 TATCCCGAATCGCCTGGCGTTTTTGCACATCCTCTGCGACGCTATTTTTGTGGAACGCAA 27245  
 Qy 61 AGCCATCAAGGAAAAATAGCCACCGGTTTAATGGCAT 96

Db 27244 AGCCATCAAGGAAAATAGCCACCGGTTTAATGGCAT 27209

Comp = Seq. 5

RESULT 1  
US-09-453-702B-242/c  
; Sequence 242, Application US/09453702B  
; Patent No. 6365723  
; GENERAL INFORMATION:  
; APPLICANT: Blattner, Frederick R.  
; Burland, Valerie  
; Perna, Nicole T.  
; Plunkett, Guy  
; Welch, Rod  
; TITLE OF INVENTION: No. 6365723e1 Sequences of E. coli O157  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Quarles & Brady  
; STREET: 1 South Pinckney Street  
; CITY: Madison  
; STATE: WI  
; COUNTRY: US  
; ZIP: 53701-2113  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette, 3.50 inch. 1.44Mb storage  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Word Perfect 8.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/453,702B  
; FILING DATE: 03-Dec-1999  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/110,955  
; FILING DATE: 04-DEC-1998  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Seay, Nicholas J.  
; REGISTRATION NUMBER: 27386  
; REFERENCE/DOCKET NUMBER: 960296.95017  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (608) 251-5000  
; TELEFAX: (608) 251-9166  
; INFORMATION FOR SEQ ID NO: 242:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31880  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; SEQUENCE DESCRIPTION: SEQ ID NO: 242:  
US-09-453-702B-242

Query Match 100.0%; Score 81; DB 4; Length 31880;  
Best Local Similarity 100.0%; Pred. No. 1.1e-21;  
Matches 81; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 CGGTTTAATGGCTTGTGTGGTAACACCGAAGCCAGCTCAATAAATTGCTGCGATGAGTT 60  
|||  
Db 27768 CGGTTTAATGGCTTGTGTGGTAACACCGAAGCCAGCTCAATAAATTGCTGCGATGAGTT 27709  
Qy 61 ACAGCTATCGAGTAAACCACC 81

Db 27708 ACAGCTATCGAGTAAACCACC 27688